STORET Import Module (SIM)

Web Version

User Guide and Reference Manual

U.S. Environmental Protection Agency
July 30, 2003

Table of Contents

Section	Page
Introducing SIM	1
Using This Guide	1
Formatting Your Data	2
Lesson I. Defining an Import Configuration	3
Lesson II. Using SIM to Migrate Data to STORET	15
Lesson III. Advanced Functions	18
Appendix A: Listing of Available STORET Data Elements for Each File Type	27

Introducing SIM

The STORET Import Module (SIM) is a Web program that helps users load data into STORET, an EPA-maintained database of ambient environmental data. STORET is a distributed system that allows individual agencies to manage their own data at a local level and to share their data with others via a national data warehouse.

To use SIM, you must perform the following tasks:

- Organize your data into delimited text files
- Establish SIM import configurations that describe the format of your text files
- Import your text files into SIM and migrate the data that pass validation into STORET.

This user guide and reference manual will

- Guide you in creating your delimited text files
- Teach you how to create a SIM import configuration
- Walk you through the process of using SIM to load data into STORET.

Note: To use SIM successfully, you need a clear understanding of STORET. Before you begin to use SIM, you should use the STORET Web Registration Application to establish all relevant metadata defaults in STORET.

Using This Guide

This guide is designed to provide you with "hands on" instruction for using SIM. Each lesson focuses on different functions included in SIM, describes the functions and associated concepts, and provides step-by-step instructions for using the tools. The step-by-step instructions are called out in shadow boxes, which are separate from the text (see box below).

1. This is an example of a box that contains instructions for using the guide with the sample data.

For assistance, contact the Central Data Exchange (CDX) Help Desk at

- EPACDX@CSC.COM
- 888-890-1995

Formatting Your Data

Before using SIM, you must organize your data into delimited text files (ASCII flat files). Most organizations use common software products, such as Excel, Access, or Lotus 1-2-3, to create data tables that can be saved and exported as delimited text files. Larger organizations may produce these files as exports from existing data management or laboratory information management systems (LIMS). SIM accepts two different file types:

1. Station Descriptions—Locations where field measurements are made or samples are collected

2. Field Measurements and Samples—Results of field measurements or sample analyses.

The data in the two file types complement each other and the order in which you migrate data files to STORET is important. Station descriptions should always be loaded before field measurements and samples. This ensures that the key data elements will be registered in STORET and are ready for results to be associated with them.

For ASCII flat files, the data elements or fields in each file type are organized in a tabular format where the column delimiter is a tab ( ), pipe (|), tilde (~), or comma (,).[1] Additionally, a subdelimiter (\) is available for columns that can accept multiple values.[2] The required and optional elements for each file type are described in Appendix A. Many of these data elements have a list of valid values that you must select from. The order in which you place the data elements in your delimited text files is not important because SIM allows you to define the order before you import your file.

Note: Do not include a header row in the delimited text files that you create, and do not use text qualifiers for text values (e.g., “”), because header rows and text qualifiers can result in errors in your import files.

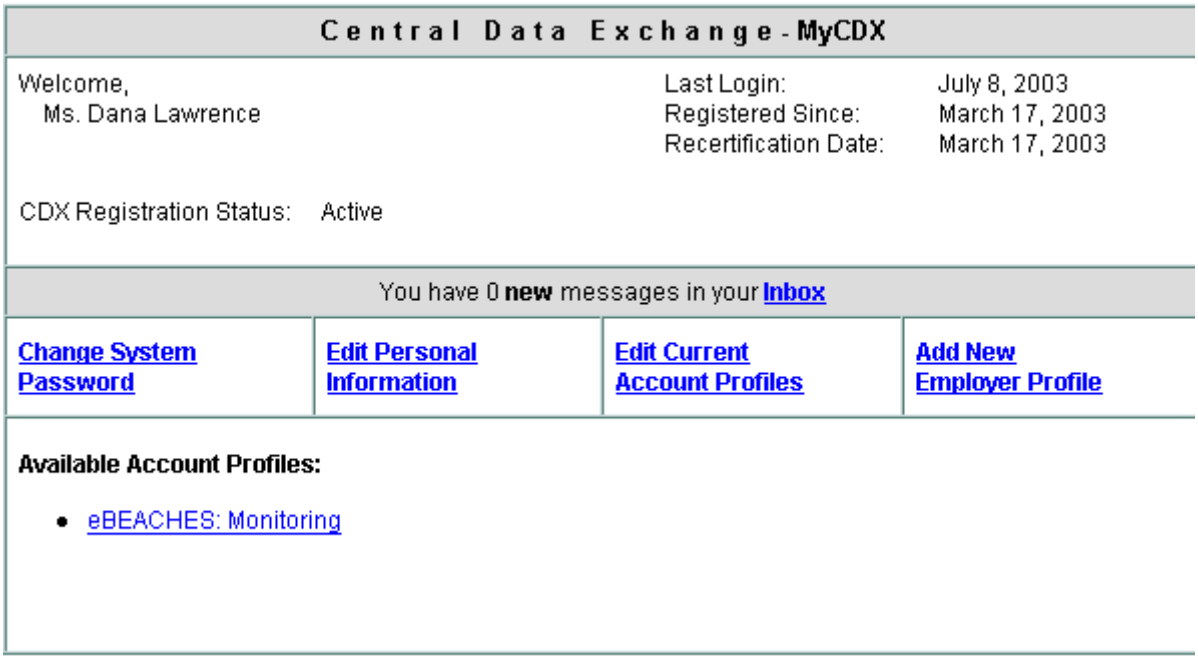
[1] If any of these characters appear in your data, you should not use that character as a delimiter because there is no way for SIM to tell which is the delimiter and which is simply part of the data.
[2] Certain fields, such as Project ID in the “Field Measurements and Non-Biological Samples” section, can accept multiple values in a single column and row of the import file. See Appendix A for a list of columns that support loading multiple values.

Lesson I. Defining an Import Configuration

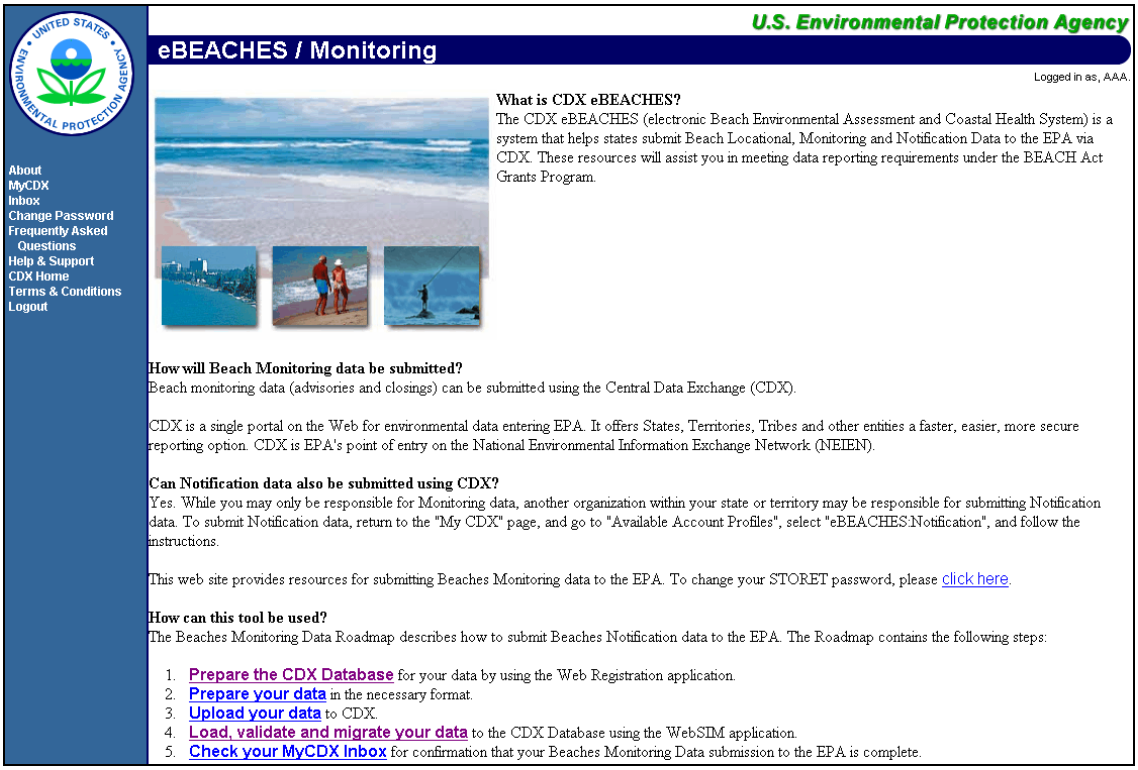
The following lesson will show you how to create a custom import configuration that will describe the format of your text files.

1. First, log in to CDX and access "MyCDX" from http://cdx.epa.gov/warning.asp. Read the warning notice and privacy statement and click the link at the bottom of the page to advance.

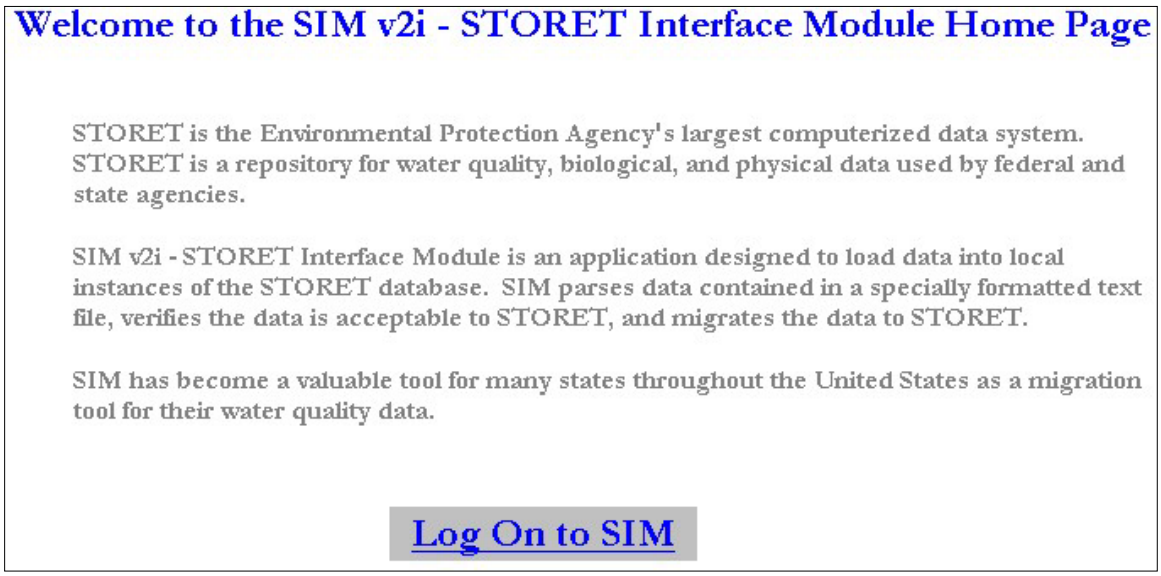
   Next, click the link for users who already have a CDX ID and password (the second choice in the bulleted list). Enter your CDX ID and password and click **LOGIN**. This will give you access to the MyCDX page. You should have access to the monitoring role to transmit beach monitoring data. If monitoring is not an option under the MyCDX Profiles, call the CDX Help Desk. Click on **eBEACHES: Monitoring** near the bottom of the page.



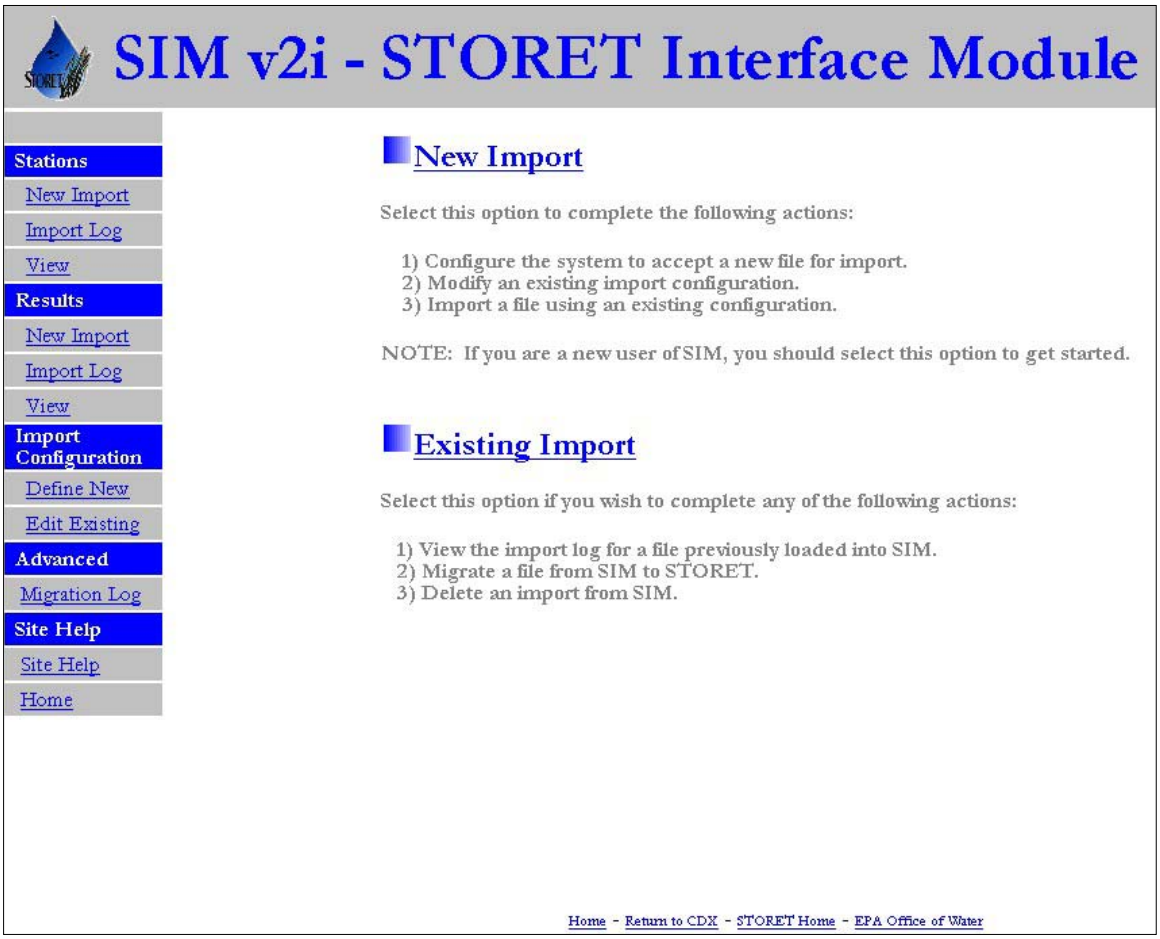
2. Click **Load, validate, and migrate your data** (item 4 near the bottom of the page).



3. Log on to SIM by clicking **WebSIM application**. On the page shown below, click **Log On to SIM** and enter the appropriate user name and password.

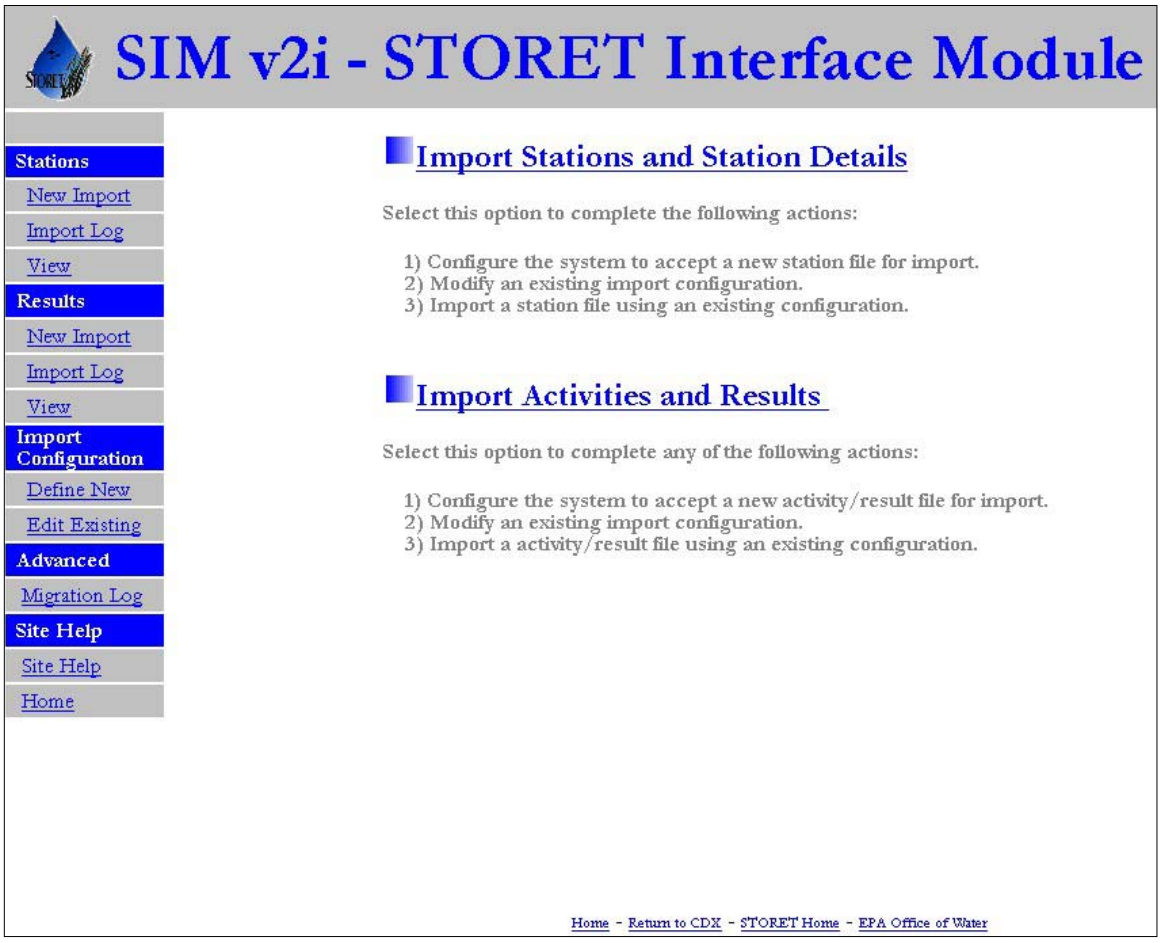


4. From this screen, you can either start a new import of data or access an existing import that you may have brought into SIM but have yet to migrate to the STORET database. For this lesson, click **New Import**.



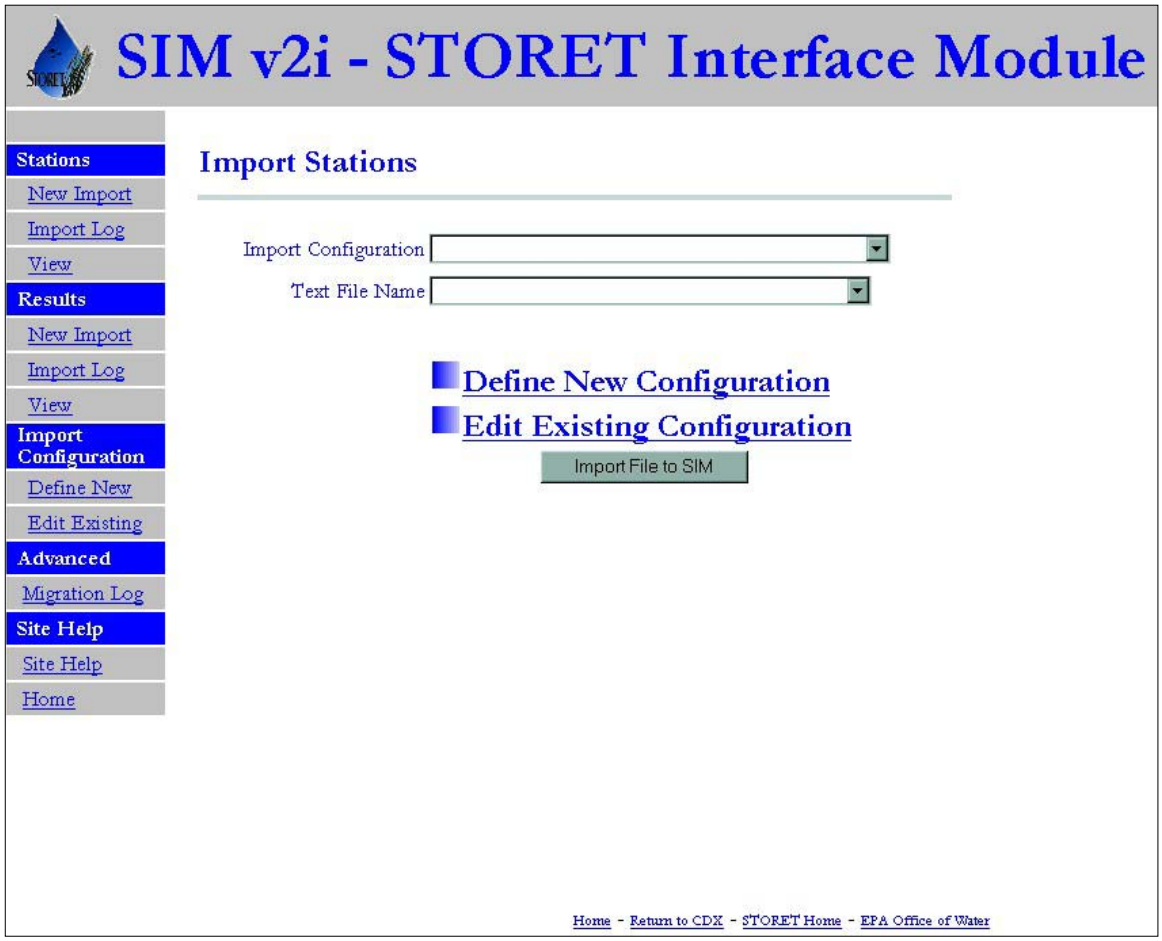
5. The next screen allows you to select the type of file that you wish to import into SIM.

   Choose **Import Stations and Station Details**.



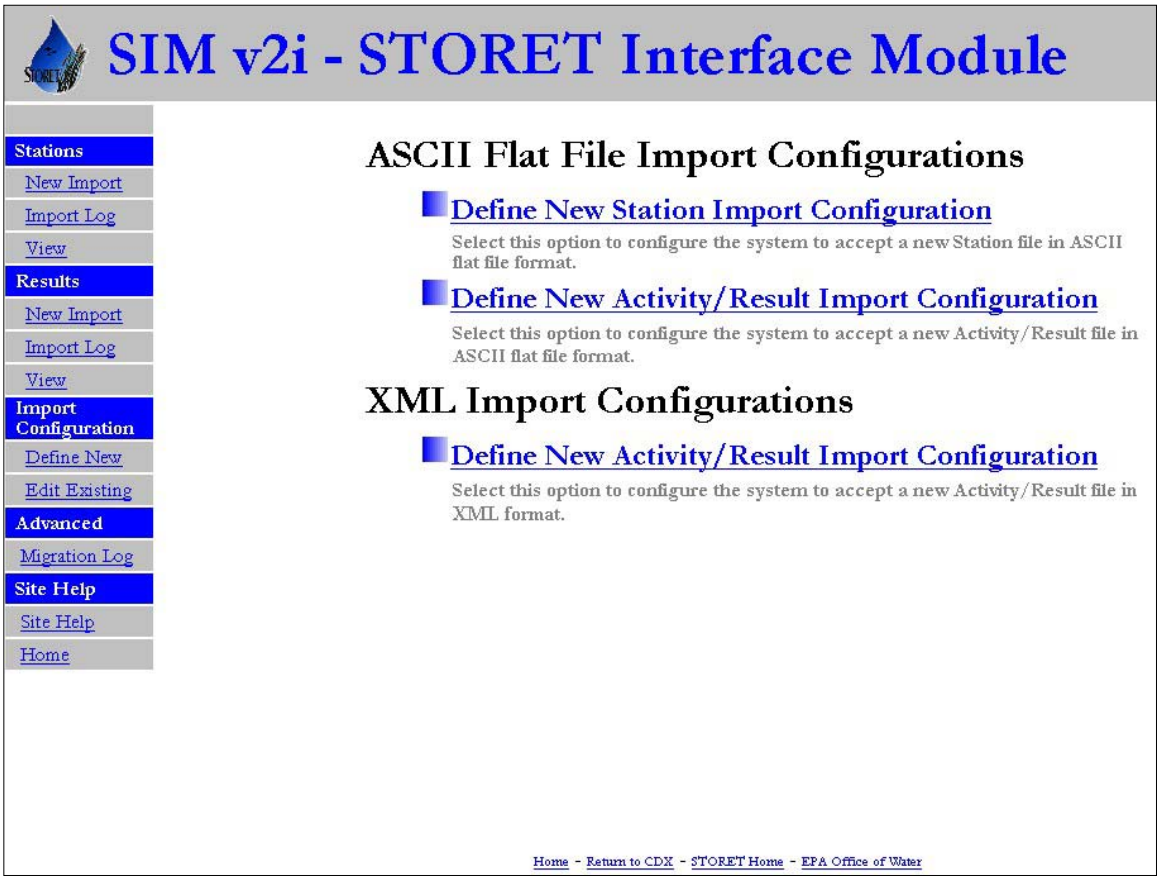
6. This menu allows you to specify the import configuration and the path to the file that you want to import.

   Click **Define New Configuration.**

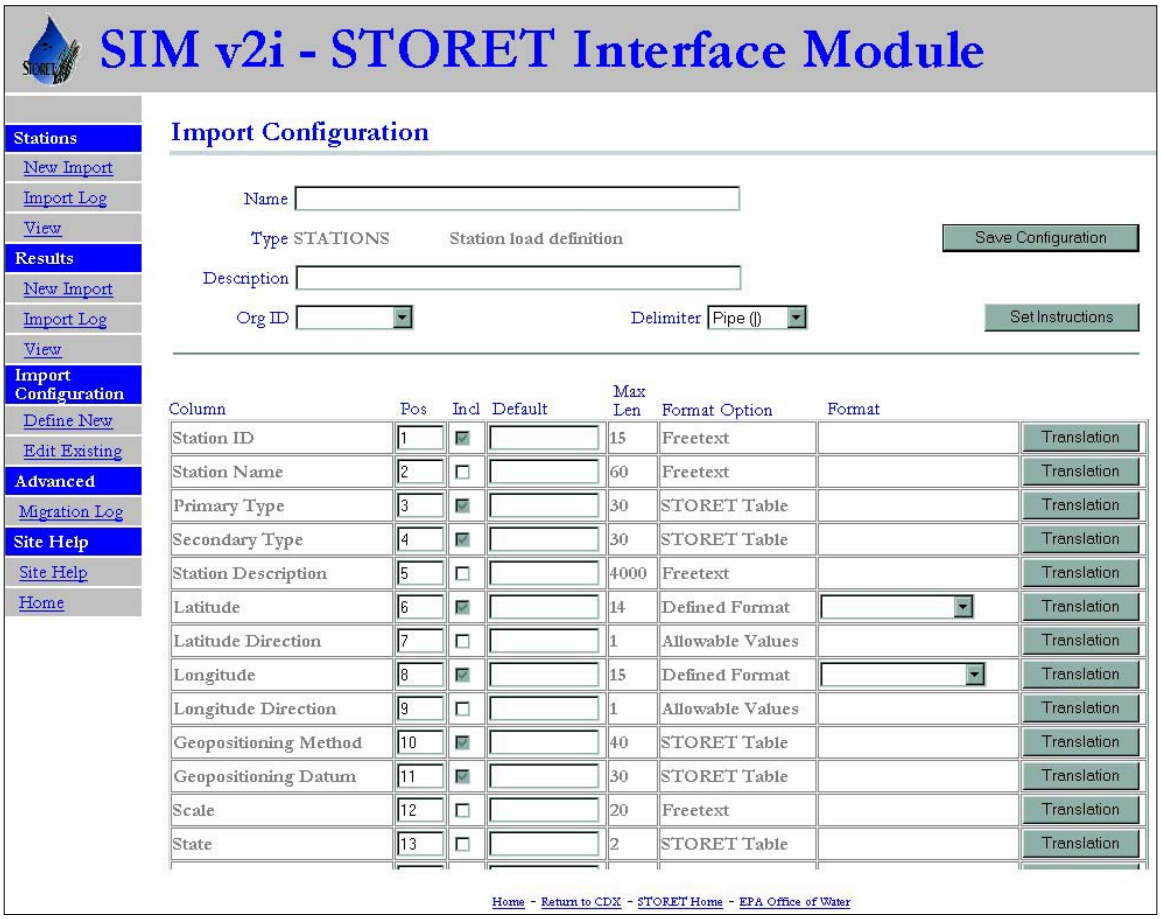


7. This menu allows you to specify the type of file that you want to import, either an ASCII flat file or an XML file. (XML files can only be imported for results.)

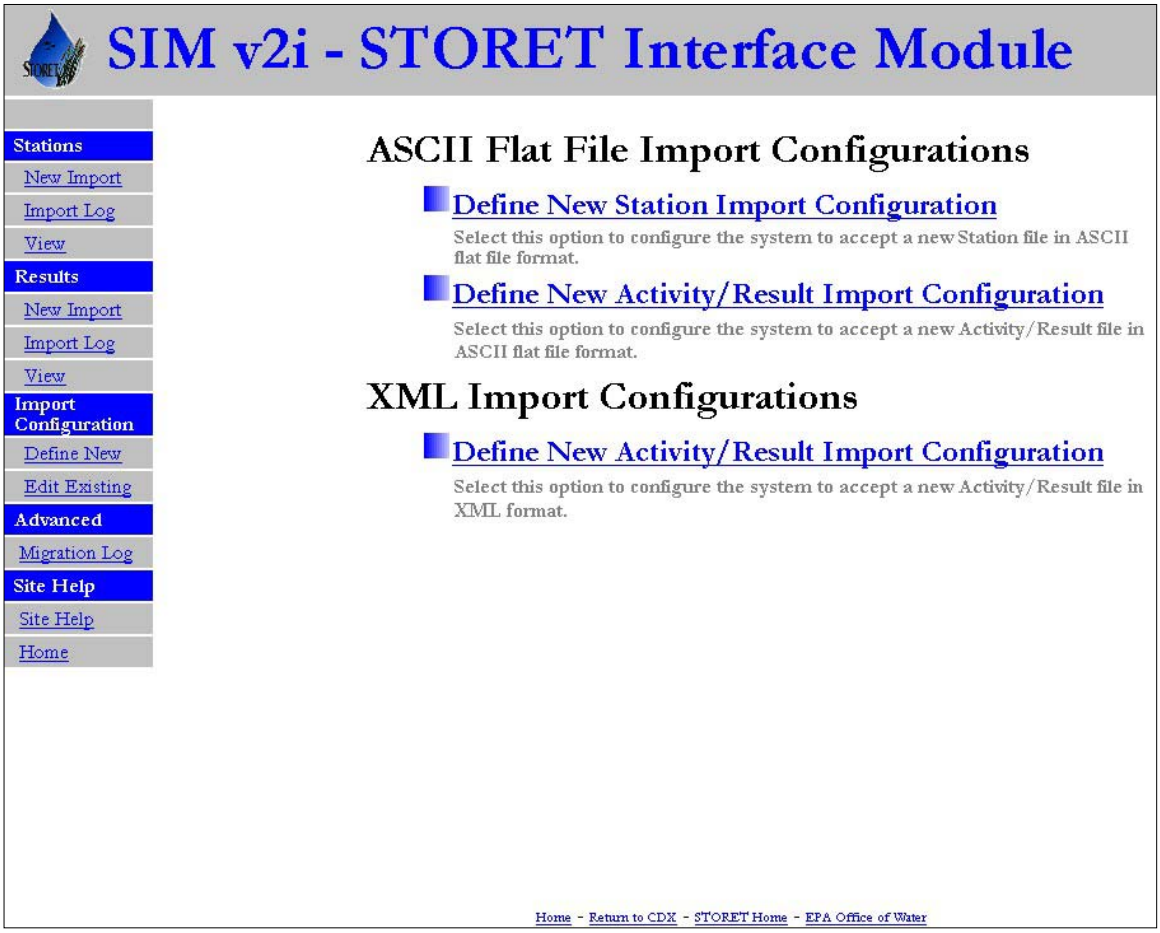
   Click **Define New Station Import Configuration** under "ASCII Flat File Import Configuration."



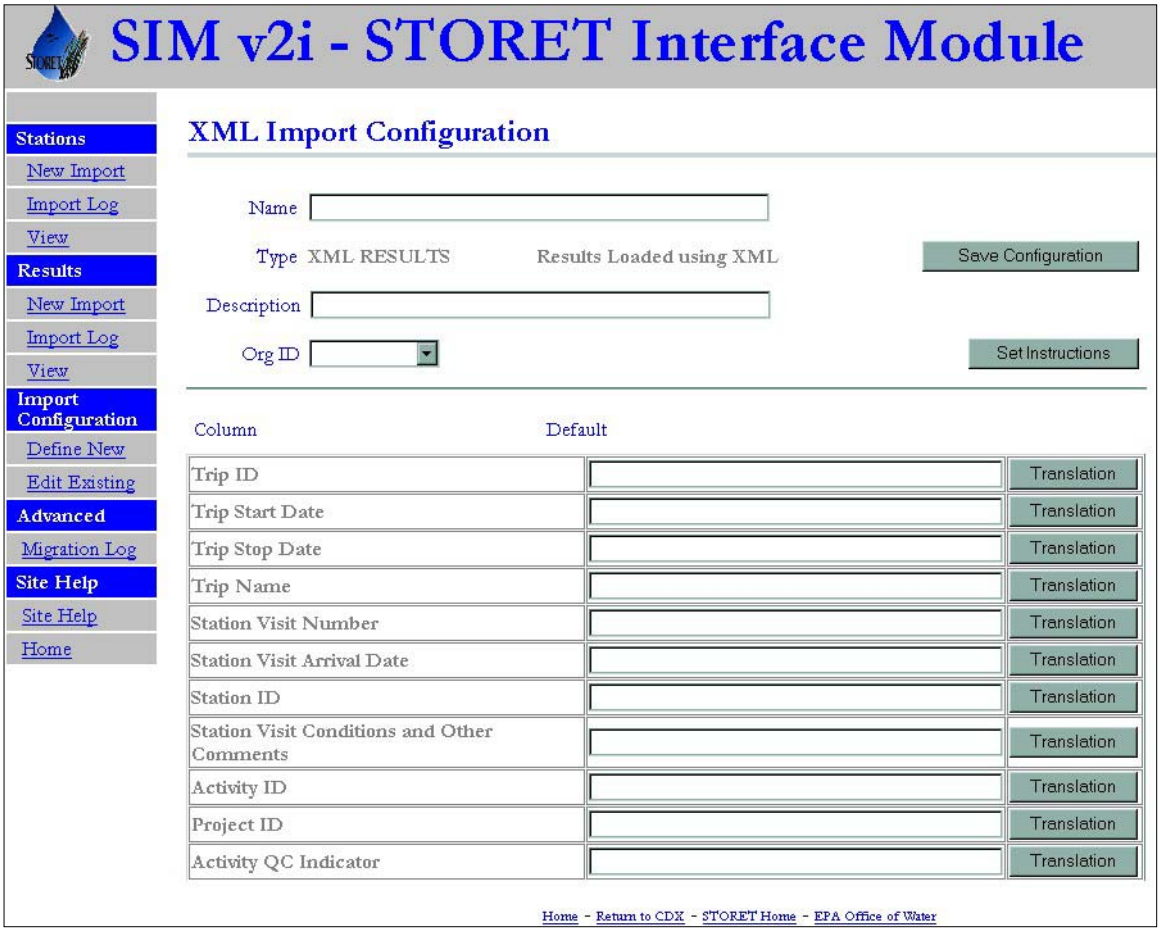
8. From the "Flat File Import Configuration" menu, you can
   - Provide a name and description for the import configuration
   - Define the organization that the data will load into
   - Indicate the delimiter that will be used in the text file
   - Save the configuration
   - Set instructions to review translations and establish general import instructions
   - Select the data elements that will be included in the text file
   - Define the position or order in which the data elements will appear in the file
   - Establish default values for the data elements in the import file that are left null or blank
   - View the format that each row in the import file must adhere to if this configuration is used
   - Specify additional formatting for individual data elements (if appropriate)
   - Establish translation parameters that change values while a file is being imported.



Click the **Back** button in your Web browser.



If you choose **Define New Activity/Result Import Configuration** under "XML Import Configurations," you can perform essentially the same functions as listed above for flat files. You do not, however, need to indicate the delimiter, because delimiters are not necessary for XML.



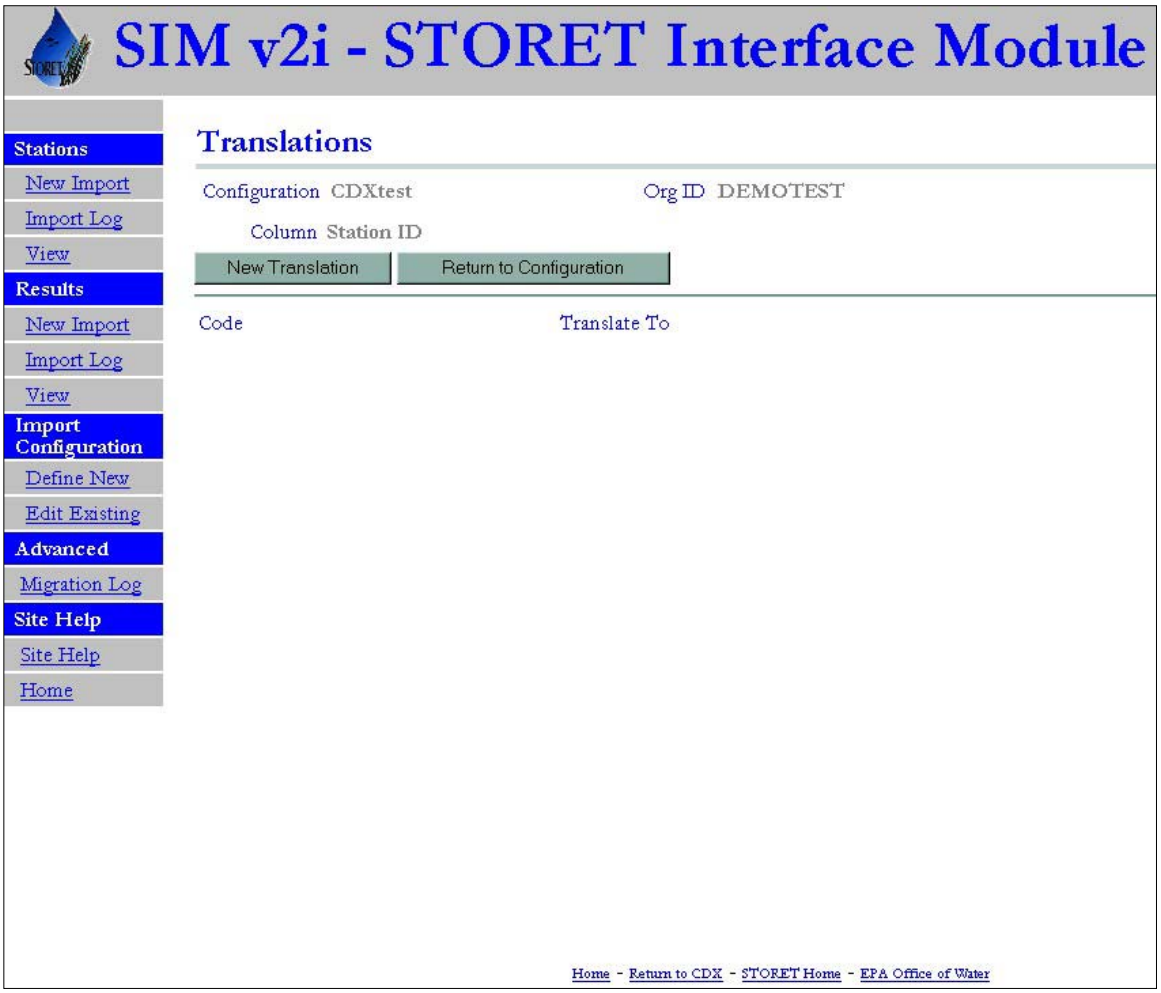
You can use a translation to convert a number or phrase in your current data to one that is accepted by STORET. If you wish to perform a translation after defining an import configuration, click the **Translate** button for the column you want to translate.

The main purpose of the XML import configuration is to set translations. You may want to establish translations to clarify your data. For example, if you have a characteristic in your data, "DO," for dissolved oxygen, you may want to translate "DO" to "Dissolved Oxygen" through the import configuration.

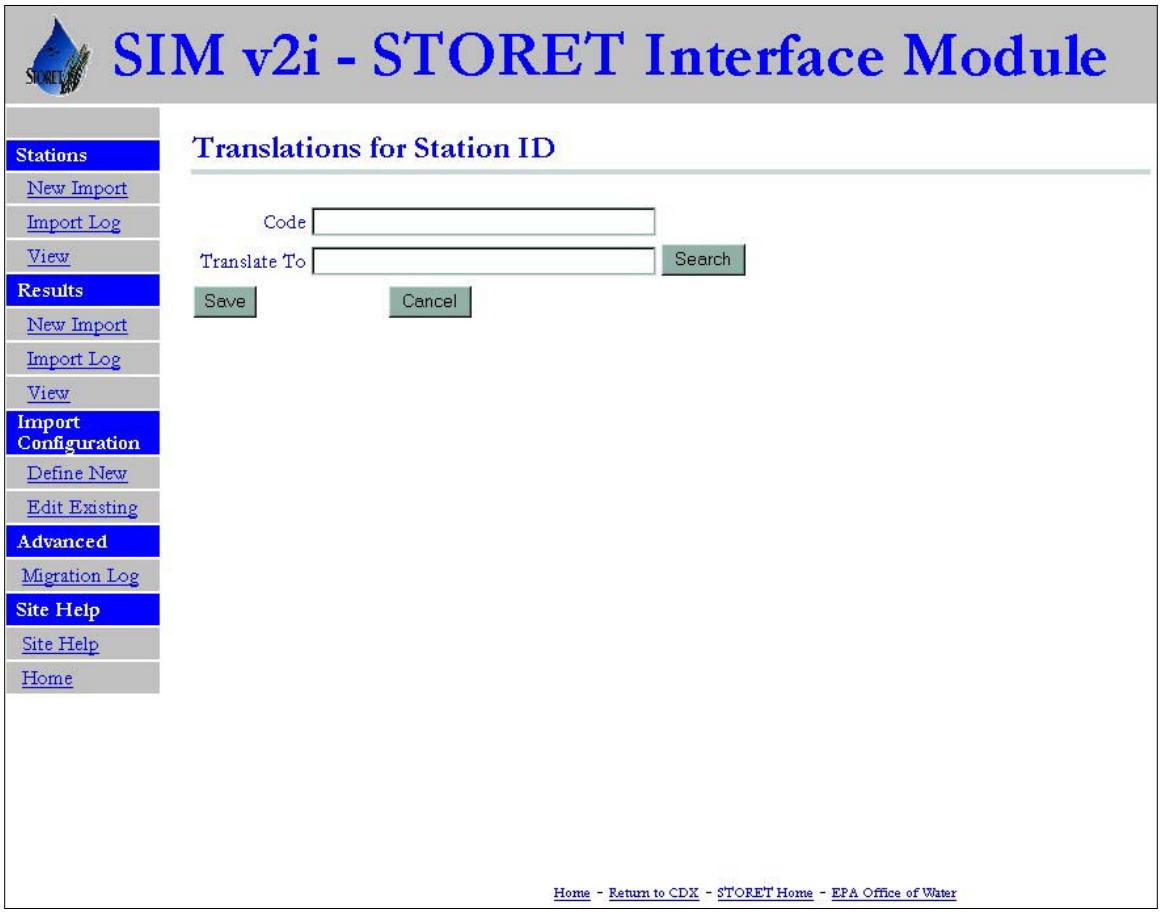
(To set up your XML schema, review the example schema on your MyCDX page. From the MyCDX page, click **eBEACHES: Monitoring**. Next, choose item 2, **Prepare your data**, to see the example.)

Provide a name and your ORG ID and click **Save Configuration**. Now you can access the Translation feature.

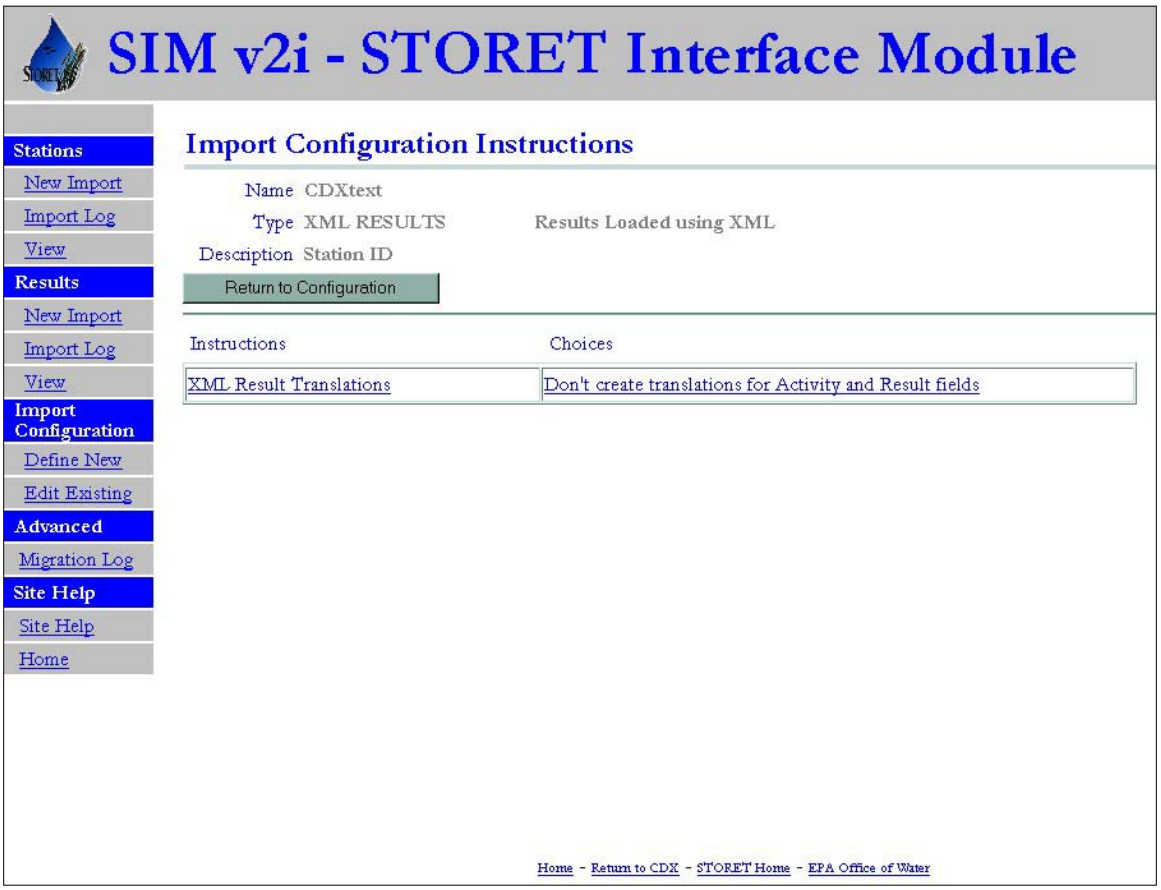
9. You can select **New Translation** or **Return to Configuration.** For this lesson, click **New Translation**.



10. You can establish a translation by defining the "code" that should be translated and the value that it should be "translated to." Click **Save.**



11. Click **Return to Configuration** and then **Set Instructions** on the Import Configuration menu.



12. You can use this menu to review translations and establish general import instructions. To change an instruction, click the name of the instruction. Press the **Back** button if you want to return to the main Import Configuration screen.

Summary

The purpose of an import configuration is to set up a pattern to translate your data file into STORET using SIM. SIM import configurations can help you set up the order of fields, format data, and perform simple data translations. Ideally, you will set up a few configurations that you will use each time future data are imported to STORET.

In Lesson I, Defining an Import Configuration, you learned how to

- Define a new import configuration
- Save a new import configuration
- Use translations to change values while a file is being imported.

Lesson II. Using SIM to Migrate Data to STORET

This lesson describes how to

- Import a file into SIM so that it can be validated against the STORET requirements
- Migrate data that pass validation into STORET.

1. Use the sidebar to select **New Import** under "Stations."

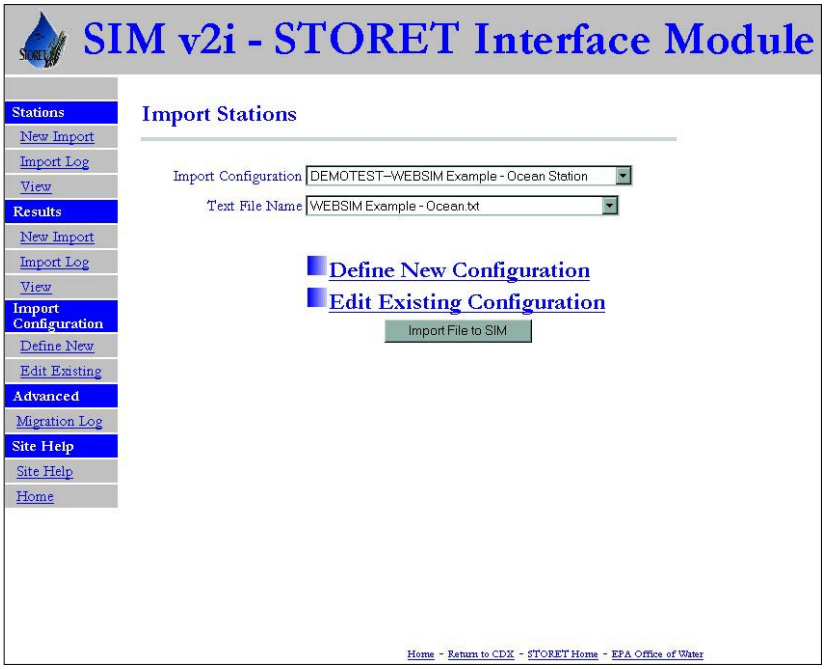
2. You must first select an existing import configuration. Click the drop-down box next to "Import Configuration" to see a list of import configurations that have been created in SIM. You can review the configuration details by clicking the **Edit Existing Import Configuration** button. In addition, you can create a new configuration by selecting **Define New Configuration**.

   Select **DEMOTEST-WEBSIM Example-Ocean Station** from the drop-down box beside "Import Configuration."

3. Then, select an appropriate text file name from the drop-down list beside "Import Configuration."

   Select **WEBSIM Example Ocean.txt**.

   Click the **Import File to SIM** button.

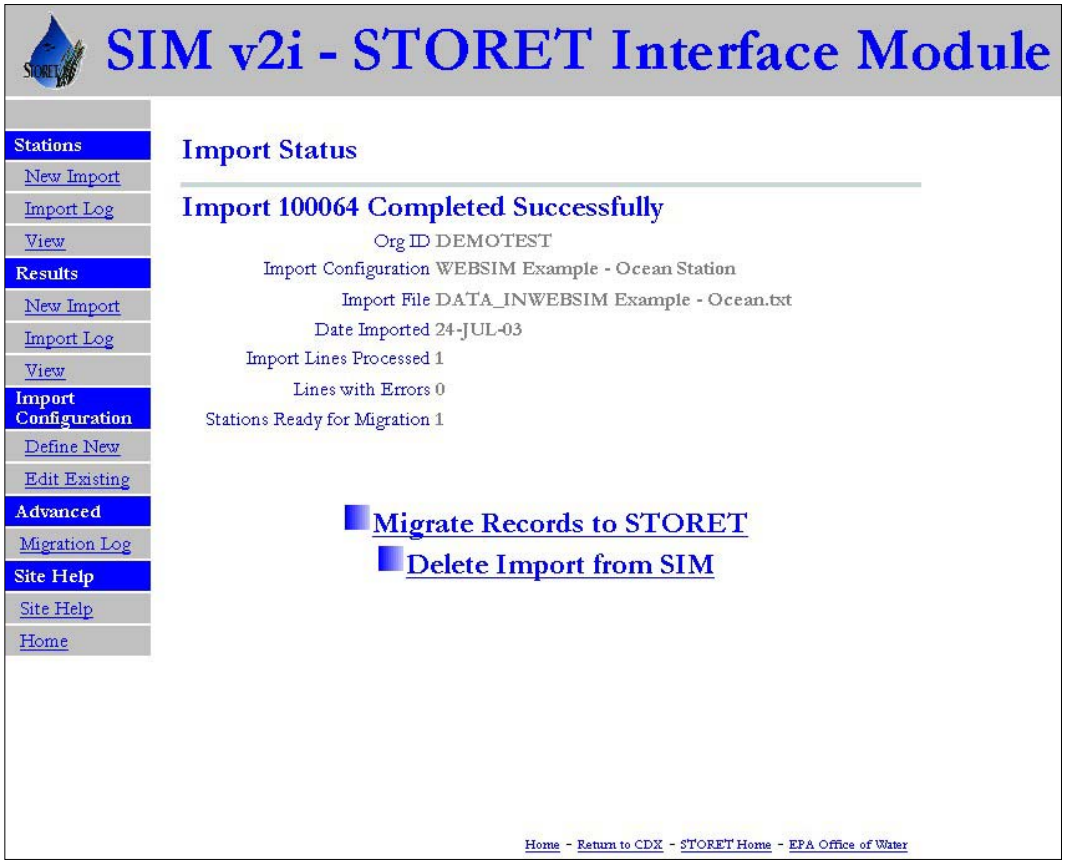


If you are importing an XML file, be sure to choose an XML import configuration.

4. SIM checks the file for errors that would prevent the data from being loaded to STORET. From the Import Status menu, you can

   - Review the number of records that that SIM imported, the number of those records that had errors, and the number of those records that are ready to be migrated to STORET
   - Migrate records to STORET
   - Delete an import from SIM. All data remain in SIM until the import is deleted. *Note: You should delete your imports once you are satisfied that your data have been accurately loaded to STORET.*

5. Click **Migrate Records to STORET** and migrate the data you imported to STORET.



Note: If your import is not successful using this example, it may be because others have already used the Ocean Station example for DEMOTEST and forgot to delete the import. Click the **View Import Errors** link. A table of import errors will appear. Clicking the **Status** button for each row will allow you to delete that import. After no imports remain, you can return to the Import Stations dialog by using the sidebar to navigate to "New Import" under "Stations" and begin the import again.

Summary

In Lesson II, using SIM to Migrate Data to STORET, you learned how to

- Import files to SIM using a predefined import configuration and text file
- Migrate data into STORET.

Lesson III. Advanced Functions

This lesson explains how to

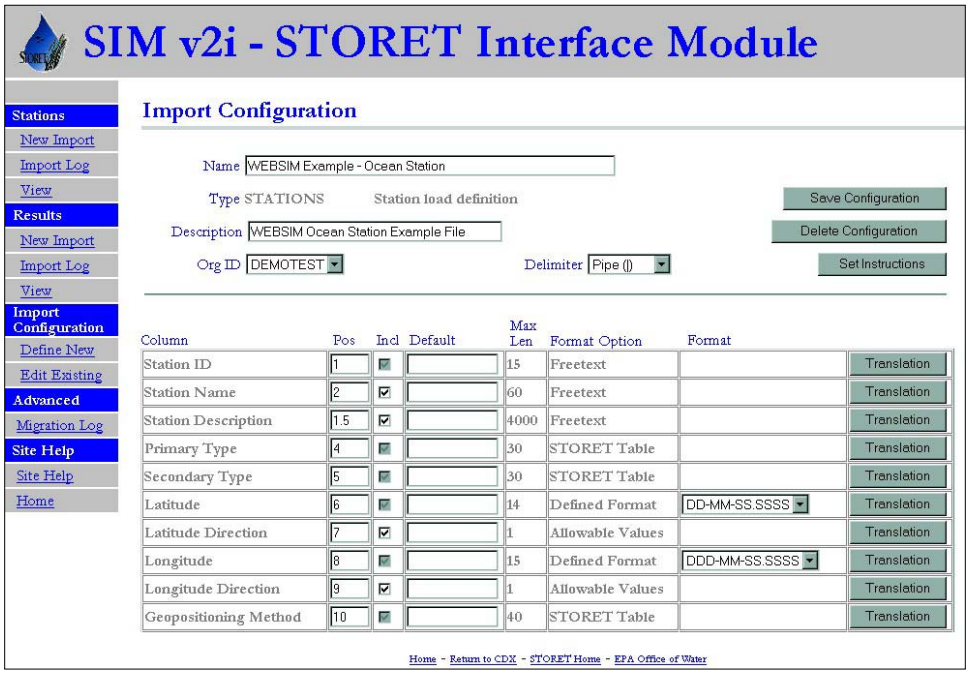
- Reorder field positions in an import configuration
- View the status of import files brought into SIM but not migrated or deleted
- Remove (undo) a migration of data from STORET
- Generate an error report to print or save
- View data that have been imported to SIM
- Test and improve system performance.

Reorder field positions in an import configuration

You can use the Import Configuration menu to reorder field or column positions.

1. From the sidebar, click **Edit Existing** under "Import Configuration." Click the **Edit** button beside the configuration you want to edit.

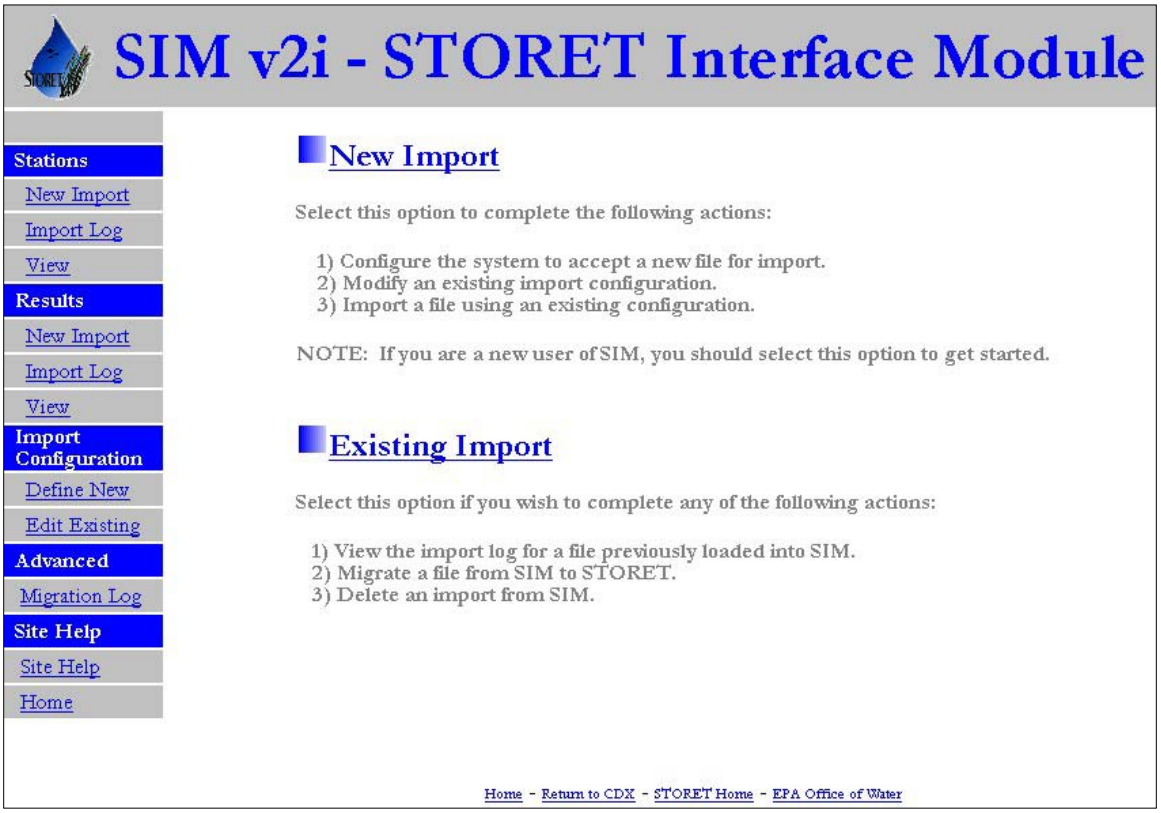
2. Suppose "Station Description" needs to be in the second position instead of the third. You can change the position number from 3 to a number between 1 and 2 (such as 1.5) and click **Save Configuration** to reorder the columns.



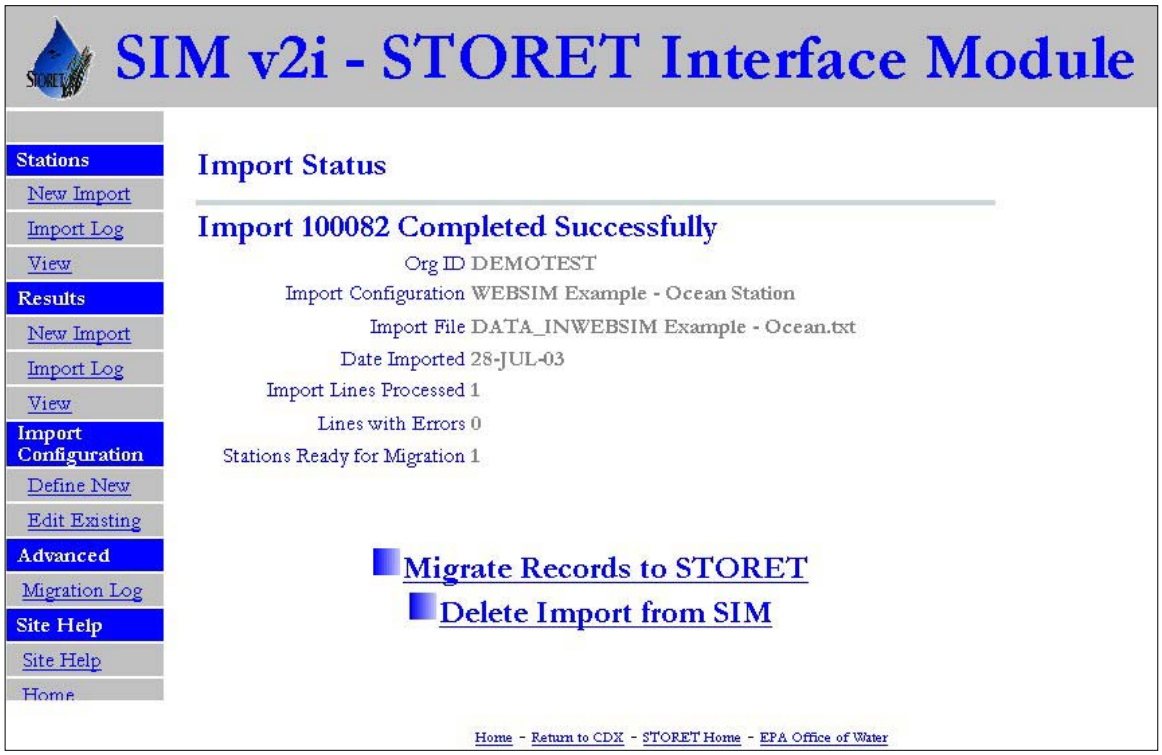
View the status of an import in SIM that was not migrated or deleted

In some cases, you may have imported files to SIM but did not migrate the data to STORET. This may have occurred because you had to stop your session or wanted to review a file. This section helps you review previous imports to determine their status. It is important to remember to delete imports after migration.

1. You can access two categories of imports from the SIM main screen: Stations and Results. To review the status of the import that was performed in Lesson II, click **Import** under "Stations" on the sidebar menu.



2. Click **Status** at the end of the row that represents the relevant import. You will return to the Import Status menu that you saw in step 3 of Lesson II.

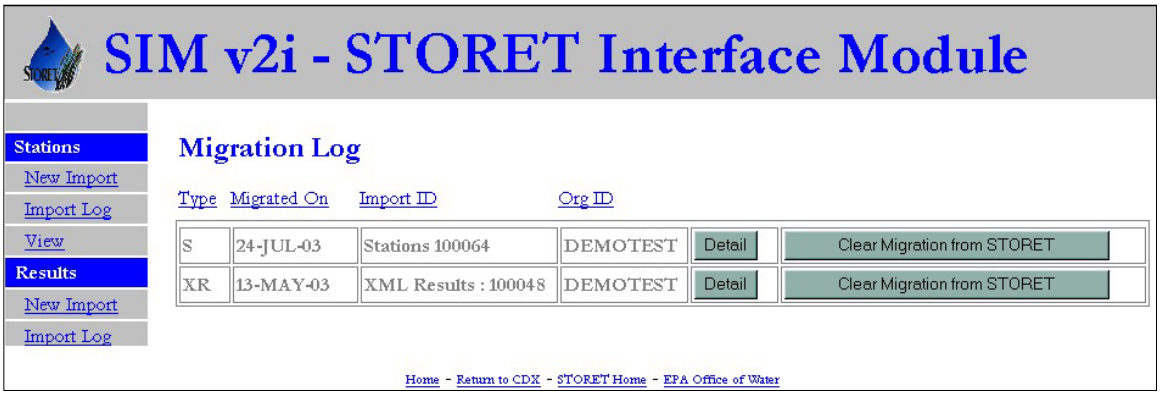


Remove a migration of data from STORET

If you migrate a file to STORET and later determine you do not want it in STORET, you can remove that data from STORET. By clearing a migration, you essentially delete the data from STORET.

1. On the sidebar, click **Migration Log** under "Advanced."

2. Click the appropriate **Clear Migration from STORET** button.

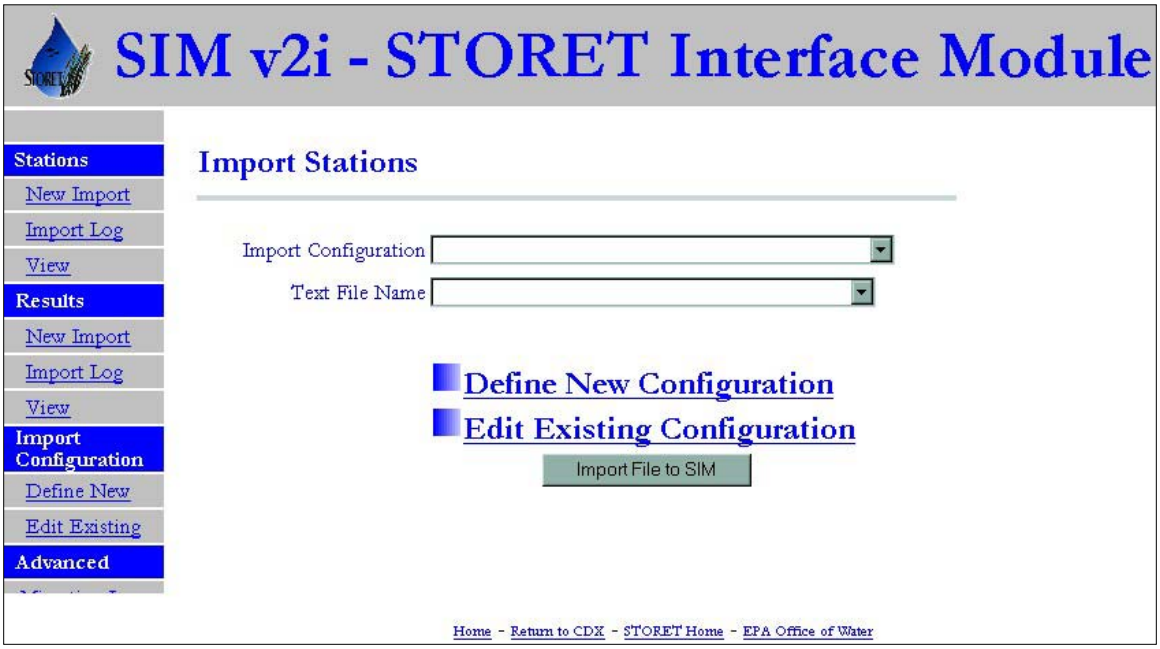


Note: SIM cannot undo a migration once the import has been deleted from SIM.

Generate an error report file to print or save

If your file is imported to SIM with errors, it is important to fix the errors and reimport the file to SIM. Reviewing an error report will help you determine which fields need to be fixed and why.

1. Click **New Import** under "Stations" on the sidebar menu.

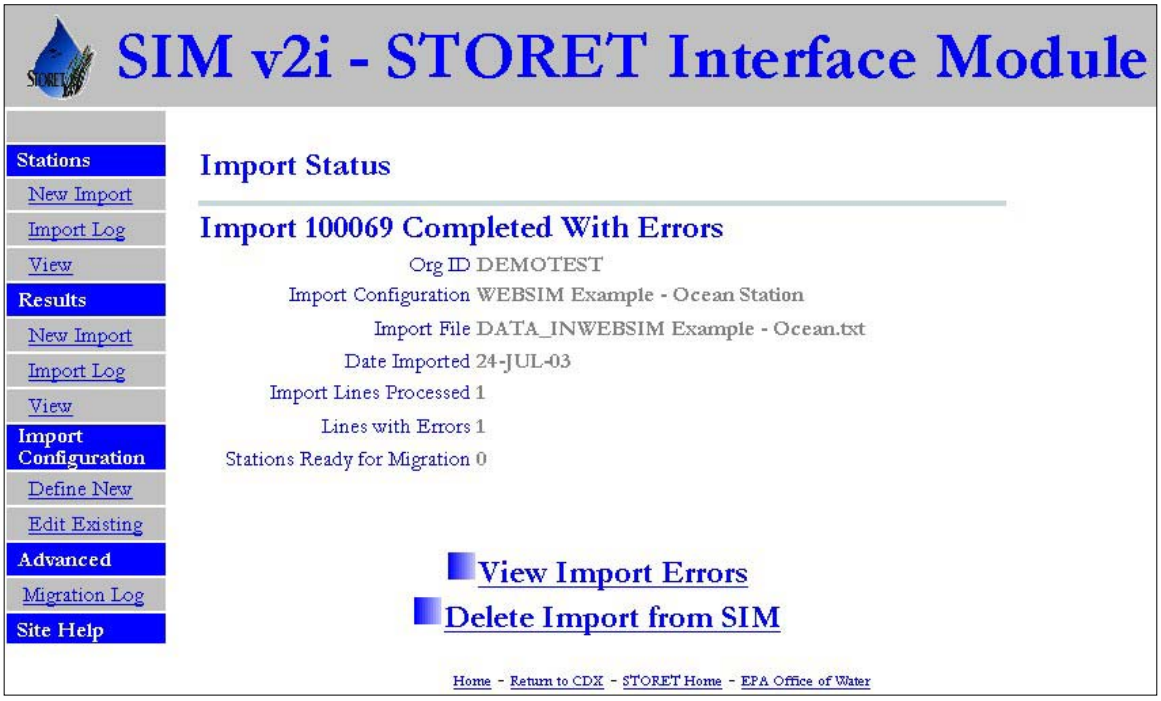


2. In order to generate an import with errors, import the same file that you imported in Lesson I.

   Select **DEMOTEST-WEBSIM Example-Ocean Station**.

   Select **WEBSIM Example Ocean.txt**.

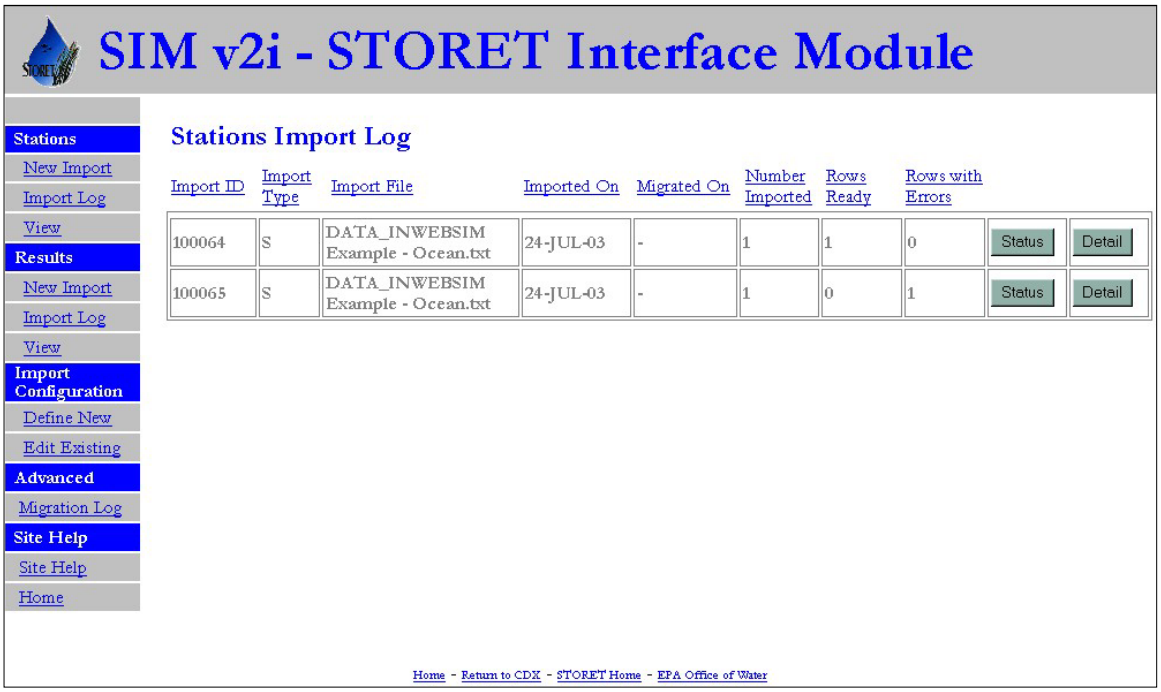
   Click **Import File to SIM.**



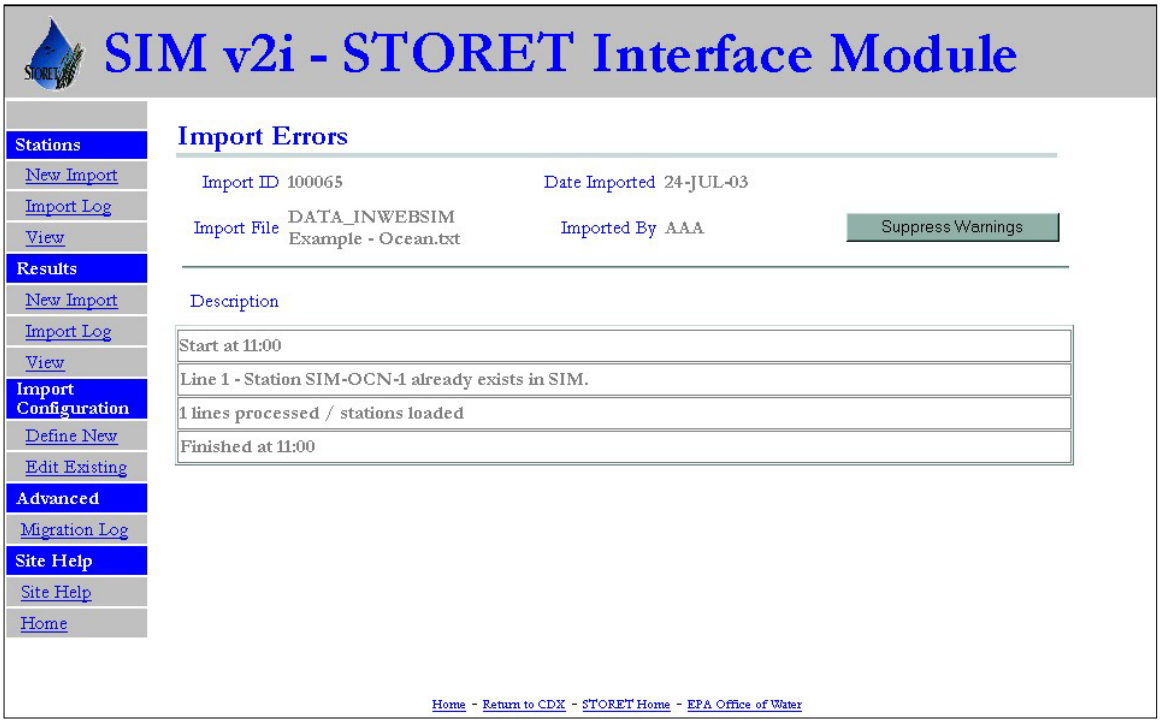
Note: If there are any errors in your file, you cannot migrate any records to STORET. In addition, if you import a file with errors to SIM, make sure to delete the import file from SIM before you import a new file. If you do not delete the import file with errors, SIM will automatically delete it when you import a new file.

3. Click **Import Log** under "Stations" on the sidebar menu.

4. Click **Detail** from the import with errors (in this case, the second import).

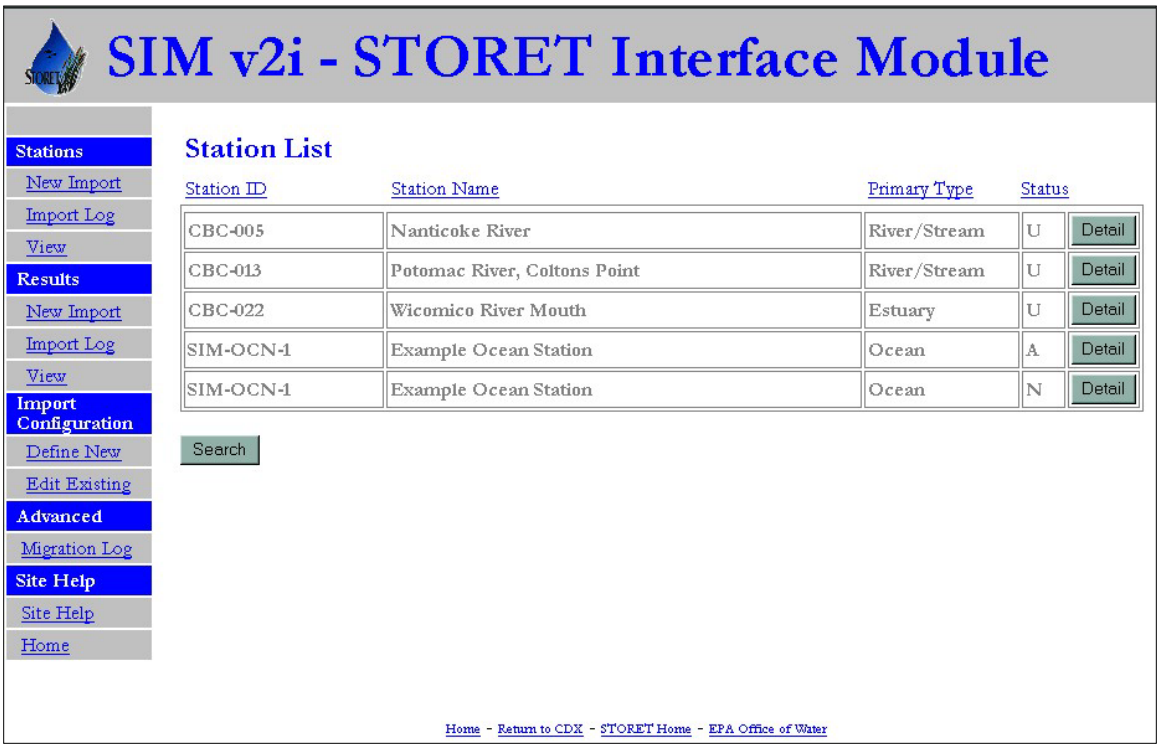


5. A report of errors that occurred during import is displayed.

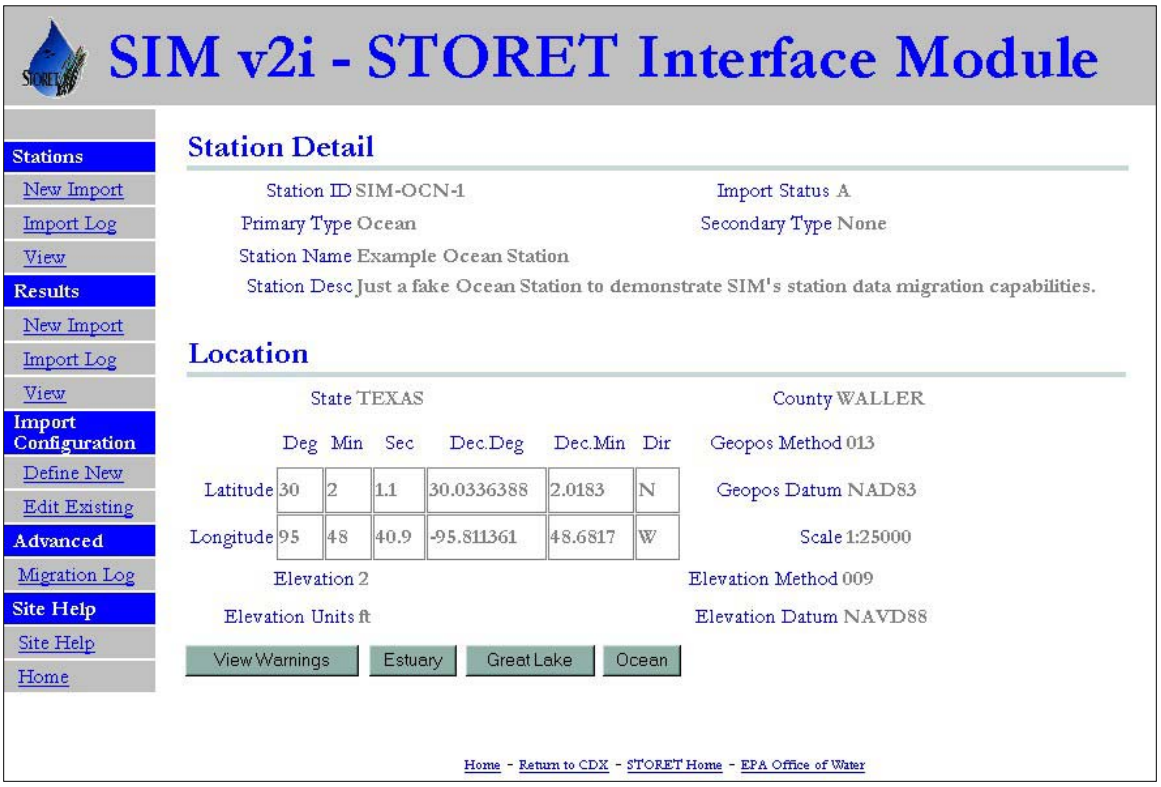


View data that have been imported to SIM

1. To view data that are in SIM, use the View capability under the Stations and Results menus. Click **View** under "Stations" from the sidebar.



2. Click **Detail** after an individual record to recall and view the data.



Test and improve system performance

- To ensure that the system runs smoothly and will not produce chronic errors, you may want to load a small sample of your data first to test for problems. This may help you avoid errors in multiple records due to an incorrect date format (or other minor error).

- Warnings slow down the import. If you limit the number of warnings that you would receive during an import (e.g., latitude direction), then the overall speed of the import will improve.

Appendix A.
Listing of Available STORET Data Elements for Each File Type

This appendix lists data elements that you can enter to describe stations, activities, and results. The tables indicate which elements are required and whether you can enter specific values for each element. These tables may be useful if you are importing flat files to SIM.

For XML import files, review the example schema on your MyCDX page to determine how to set up your XML schema. From the MyCDX page, click **eBEACHES: Monitoring**. Next, choose item 2, **Prepare your data**, to see the example.

Table A-1. Station Descriptions

Data Element	Required	Allowed Values
Station ID	Yes	
Station Name	Yes	
Primary Type	Yes	Yes
Secondary Type	Yes	Yes
Station Description		
Latitude	Yes	
Latitude Direction	Defaults to “N”	Yes
Longitude	Yes	
Longitude Direction	Defaults to “W”	Yes
Geopositioning Datum	Yes	Yes
Geopositioning Method	Yes	Yes
Scale	Conditional	
State	Yes	
County	Yes	
Elevation		
Elevation Units	Conditional	Yes
Elevation Method	Conditional	
Elevation Datum	Conditional	
Ocean Name	Conditional	Yes
Shore Relation	Conditional	Yes
Additional Ocean Name		
Primary Estuary	Conditional	
Secondary Estuary		
Other Estuary		
Great Lake	Conditional	Yes
Additional Great Lake Name		

Table A-2. Activities and Results

Data Element	Required	Allowed Values	Multiple Allowed
Trip ID	Yes	Must Exist in STORET	
Trip Start Date			
Trip Stop Date			
Trip Name			
Station ID	Yes	Must Exist in STORET	
Additional Location Information			
Station Visit Number	Yes		
Station Visit Arrival Date			
Visit Comments			
Project ID	Yes	Must Exist in STORET	Yes. Separate with "\"
Activity ID	Yes		
Medium	Yes	Yes	
Activity Type	Yes	Yes	
Activity Category	Conditional	Yes	
QC Indicator		Yes	
Sample Matrix		Yes	
Replicate Number	Conditional		
Activity Start Date	Yes		
Activity Start Time			
Activity Start Time Zone	Conditional	Yes	
Activity End Date			
Activity End Time			
Activity End Time Zone	Conditional	Yes	
Depth to Activity			
Depth to Activity Units	Conditional	Yes	
Relative Depth		Yes	
Depth Measured From			
Activity Comments			
Sample Collection Procedure ID	Conditional	Must Exist in STORET	
Gear ID	Conditional	Yes	
Gear Configuration ID		Must Exist in STORET	
Gear Deployment Comments			
Sample Preservation, Transport & Storage ID		Must Exist in STORET	
Sample Transport and Storage Comments			
Detection Condition		Yes	
Characteristic Group ID	Conditional	Must Exist in STORET	
Characteristic Row ID	Conditional	Must Exist in STORET	
Characteristic Name	Conditional	Yes	
Result Value	Conditional	Yes*	
Result Value Units	Conditional	Yes	
Result Status	Defaults to "F"	Yes	
Sample Fraction	Conditional	Yes	
Statistic Type		Yes	
Value Type	Defaults to "Actual"	Yes	
Duration Basis		Yes	
Temperature Basis		Yes	
Weight basis		Yes	
Result Comment			
Laboratory ID		Must Exist in STORET	
Field/Lab Procedure	Conditional	Must Exist in STORET	
Field/Lab Procedure Source	Conditional	Yes	
Laboratory Certified		Yes	
Laboratory Batch ID			
Analysis Date			
Analysis Time			
Analysis Time Zone	Conditional	Yes	
Detection Limit			
Detection Limit Unit	Conditional	Yes	
Detection Limit Comment			